Vol. 1, No. 1, 2023
ISSN: 2987-3932
https://pubs.ast-ptm.or.id/index.php/eat

Identification of Javanese Alphabet Handwritting by using Histogram Chain Code

Arief Budiman [a,1,*], Abdul Fadlil [b,2], Rusydi Umar [a,3], Adhy Kurnia Triatmaja [a,4], Muhammad Kunta Biddinika [c,5]

[a] Graduate Program of Informatics, Faculty of Industrial Technology, Universitas Ahmad Dahlan, Yogyakarta, 55166, Indonesia
[b] Electrical Engineering Department, Faculty of Industrial Technology, Universitas Ahmad Dahlan, Bantul, 55166, Indonesia
[c] Vocational Study in Electrical Engineering (PVTE), Faculty of Teacher Training and Education, Universitas Ahmad Dahlan, Yogyakarta 55161, Indonesia
[1] arief1708048029@webmail.uad.ac.id *; [2] fadlil@mti.uad.ac.id; [3] rusydi@mti.uad.ac.id; [4] adhy.triatmaja@pvte.uad.ac.id;
[5] muhammad.kunta@mti.uad.ac.id
* Corresponding Author

ABSTRACT

One of the wealth of the Indonesian nation is the number of tribes with their own language and script. One of the scripts that has existed for a long time before the independence of the Indonesian state is the Javanese script, with the use of Latin script which is used by almost every aspect of life, whether official activities or daily use, the use of traditional scripts, especially Javanese script, is increasingly rare. making it easier to learn Javanese script requires learning media with the ability to recognize Javanese characters. In this study, pre-processing is used, especially feature extraction using the Histogram Chain Code (HCC) method and classification using an artificial neural network using the Multi Layers Perceptron method. This study compares four research models by setting the number of HCC feature extraction parameters obtained from one intact image and 3 divided images of 4, 9 and 16 parts respectively so that the number of parameters of each HCC model is 8, 32, 72 and 128 parameters. characteristic. The training and testing process with the Multi Layers Perceptron method uses 2000 Javanese handwritten image data which is divided into 80%, namely 1600 images for the training process and 400 images for the testing process. This study resulted in different accuracy, namely 57%, 78%, 83% and 76%. The best accuracy is obtained from the HCC model with 72 parameters and the image is divided into 9 parts.

KEYWORDS
chain code
histogram chain code multi layer perceptron
artificial neural network
handwriting character recognition
java character

1. Introduction

One of the wealth of the Indonesian nation is the number of tribes with their own language and script. One of the scripts that has existed for a long time before the independence of the Indonesian state is the Javanese script, with the use of Latin script which is used by almost every aspect of life, whether official activities or daily use, the use of traditional scripts, especially Javanese script, is increasingly rare.

One of the efforts to preserve the nation's wealth is to introduce the culture to the early generations through technology, especially computers so that learning becomes more efficient and effective [1-5].

The introduction of Javanese script to the Indonesian population, especially the Javanese population is carried out through the world of education by including it as a local content subject, especially in the Special Region of Yogyakarta, Central Java, and East Java [2]. One of the efforts to preserve Javanese script is done by utilizing information technology, namely through handwritten character recognition technology in Javanese script.

Various methods have been developed to introduce Javanese characters. Isnaini [3] through research "Development of Flash-Based Javanese Interactive Learning Media in Javanese Language Subjects" resulted in the conclusion that Flash-Based Interactive Learning Media of Javanese Script is effective in increasing understanding in learning Javanese Script. Concerned [4] through the research "Development of Javanese Script Interactive Multimedia for Javanese Language Learning for Class V Students of SDN

Vol. 1, No. 1, 2023
ISSN: 2987-3932
https://pubs.ast-ptm.or.id/index.php/eat

Sabdodadi Keyongan Bantul, Indonesia" resulted in the conclusion that the Javanese Script Interactive Multimedia was able to facilitate students' understanding of the material for reading Javanese words using pairs.

The introduction of Javanese script handwriting characters is needed to make Javanese script learning media more interactive because it allows Javanese script learning media to provide direct feedback on students' understanding of Javanese script.

Various methods have been developed from pre-processing, feature extraction to classification of Javanese handwritten images. The Freeman Chain Code method [5] is one of the best methods used for identifying handwritten characters, several researchers have developed this method to improve the ability to extract features from handwriting such as Differential Chain Code [6], Histogram Chain Code, Vertex Chain Code [7], and several other Modified Chain Codes.

The application of the Histogram Chain Code has been used by Arora [8] to recognize Devnagari characters with an accuracy of 89%, Qian [9] has also conducted pattern recognition experiments with Number objects with an accuracy of 97.54%. In this study, the application of HCC Feature Extraction was carried out on the Ngagena Javanese script handwritten image object with the Multi Layers Perceptron classification method, by making several HCC models by setting the number of feature parameters, the number of parameters would be obtained which resulted in the best classification accuracy of several models researched.

2. Method

2.1. Research Steps

This study aims to produce a Javanese script writing identification system using Multilayer Percetron. In order to achieve the research objective, systematic steps are needed as shown in Fig. 1.

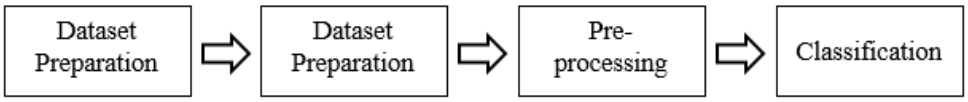

Fig. 1. Research steps

2.2. Dataset Preparation

The object of this research is the image of handwritten Javanese script with 100 data per letter in the Nglagena Javanese script which is divided into 80 training data and 20 test data. The Javanese script data set used uses a dataset created by phiard version 10 with a public license, the data set acquisition methodology is real human writing. The Javanese script handwritten image files are separated into one file for each Javanese script letter, so there are a total of 2000 Javanese script image files, with dimensions of 224x224 pixels. Some examples of datasets used in this study can be seen in Fig. 2.

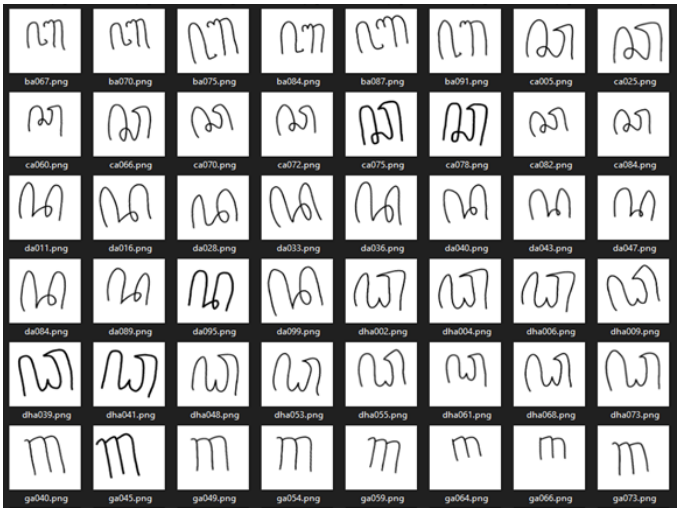

Fig. 2. Dataset files

2.3. Pre-processing

The image obtained in the dataset preparation stage is in the form of an image file in Portable Network Graphic (PNG) format with 3 RGB colour layers, so it is necessary to convert it to an image in a binary format so that edge detection operations can be carried out in preparation for the feature extraction process.

The display of the original image file is in PNG format even though it is black but because it consists of 3 colour layers it needs to be converted to binary format so that it is ready for the next process. The change from a black-and-white PNG file to a binary format is shown in Fig. 3.

2.4. Chain Code Histogram Feature Extraction

Freeman in 1961used to present digital curves and Freeman code. The Frerman Chain Code algorithm aims to represent the contours including the pixels of an object that are interconnected with the direction of the wind direction. Chain code is a simple way to present an image (Hadi, Budi, and Ramadhani, 2015). The way the chain code algorithm works is by giving a rotation mark that is adjusted to the direction of the wind that you want to use. The final result of the chain code is a feature vector that contains information on the chain code sequence forming the object (Levina and Armanto, 2015).The image obtained in the dataset preparation stage is in the form of an image file in Portable Network Graphic (PNG) format with 3 RGB color layers, so it is necessary to convert it to an image in a binary format so that edge detection operations can be carried out in preparation for the feature extraction process.

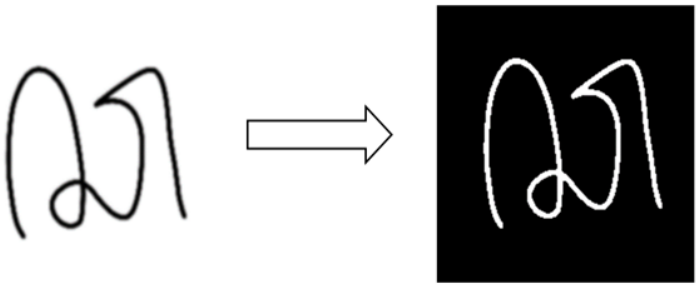

Fig. 3. Convert PNG file format to binary format

One method of feature extraction based on Freeman Chain Code is the Histogram Chain Code (HCC) formed by calculating the direction frequency of the chain code vector of the object. So that we get 8 characteristic parameters of the object. This study aims to determine the highest classification accuracy by examining 4 HCC models resulting from variations in intact images which are divided into 1, 4, 9, and 16 parts as shown in Fig. 4. The complete image is formed by cropping parts of the Javanese script, and resizing it. so they have the same size, which is 90 x 150 pixels.

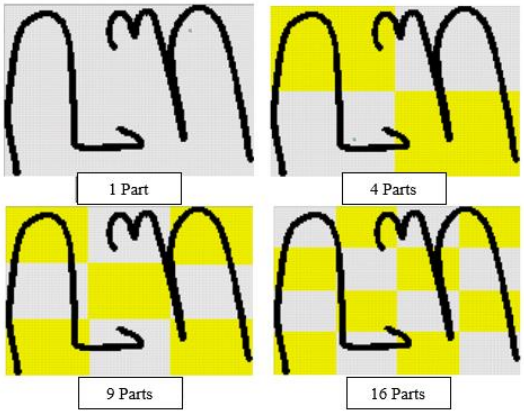

Fig. 4. Cutting image objects into 4, 9, and 16 parts

HCC feature extraction for a whole image has 8 parameters, while for a divided image, the number of parameters is obtained from the number of parts multiplied by 8. In this study, 4 HCC models were made obtained from intact images and cropped images with the number of parameters for each model shown in Table 1. The HCC Model A has an image shape of full image and Model B has an image cropped 2x2, 3x3, and 4x4 for Model C, and Model D accordingly.

Table 1. HCC model table

HCC Model	Image Shape	Number of Parts	Number of HCC Parameter
HCC Model A	Full image	1	8
HCC Model B	2x2 cropped image	4	32
HCC Model C	3x3 cropped image	9	72
HCC Model D	4x4 cropped image	16	128

2.5. Classification of Multi Layers Perceptron (MLP)

The results of feature extraction are then processed in the classification stage using the Multi Layers Perceptron (MLP) method as shown in Fig. 5.

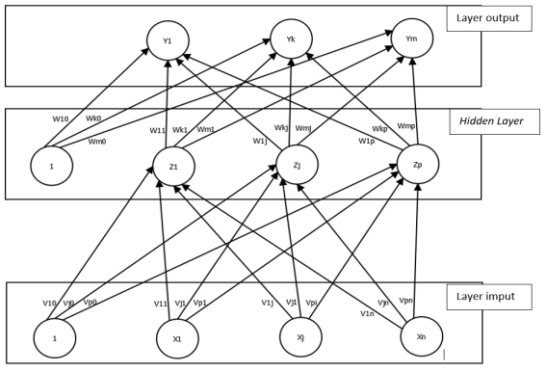

Fig. 5. Method of multi layers perceptron (MLP)

The MLP stage consists of 2 stages, namely 1) training, and 2) testing.

- MLP Training

The training phase uses 80 image data for each letter so that the total image used for the training phase is 1600 image data files

- MLP Testing

The testing phase uses 20 image data for each letter so that the total image used for the training phase is 400 image data files

The testing stages produce accuracy values by comparing the results of the test image classification with the correct classification data. The overall accuracy of the model is known simply from the number of suitable predictions divided by the total number of predictions made [10], in the form of a formula as follows:

$$Accuracy\ (\%) = \frac{Correct\ Prediction}{Number\ of\ Predictions} x\ 100\%$$

3. Results and Discussion

This research produces some data obtained by doing 4 repetitions of the research stages, namely the stages of feature extraction and classification. This repetition is in accordance with the design of the 4 HCC models so that it will produce different data.

3.1. Results of HCC Feature Extraction Method Model A

This stage is carried out by extracting the HCC features from the intact image so as to obtain a number of 8 HCC feature parameters, then the training and testing process is carried out using 1600 training image data and 400 test image data, respectively. The test results are tabulated in Table 2 with 20 Nglagena Javanese script classes.

Table 2. Test Results of HCC model A

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	LA	NA	JA	RA	KA	DA	TA	DA	WA	TA	PA	DHA	JA	YA	BA	MA	GA	NYA	THA	KA
xx002	YA	NYA	CA	RA	LA	DA	TA	DA	WA	TA	PA	DHA	JA	YA	HA	JA	JA	BA	TA	CA
xx003	LA	NYA	THA	RA	KA	DA	TA	SA	WA	YA	NA	CA	MA	YA	HA	JA	PA	BA	CA	NGA
xx004	LA	SA	DHA	RA	KA	DA	TA	SA	CA	SA	NYA	WA	WA	YA	NYA	CA	DA	NYA	JA	NGA
xx005	TA	NA	CA	RA	KA	DA	TA	DHA	CA	HA	NA	CA	JA	TA	HA	MA	GA	NYA	THA	NGA
xx006	LA	SA	JA	RA	HA	SA	SA	PA	JA	LA	PA	DHA	JA	YA	NGA	MA	GA	PA	NGA	NGA
xx007	HA	NYA	MA	RA	SA	DA	KA	KA	JA	LA	NYA	DHA	JA	YA	NYA	MA	GA	NGA	THA	NGA
xx008	HA	DA	BA	RA	KA	PA	HA	SA	WA	HA	NA	HA	JA	YA	NGA	WA	GA	BA	THA	RA
xx009	NYA	PA	CA	RA	KA	NA	TA	KA	WA	TA	PA	HA	JA	YA	NGA	WA	GA	NGA	THA	NGA
xx010	HA	CA	CA	RA	LA	DA	TA	SA	WA	KA	PA	HA	JA	YA	NYA	MA	GA	NGA	THA	NGA
xx011	SA	NA	WA	RA	TA	PA	HA	NA	WA	HA	PA	CA	NA	YA	NYA	MA	GA	NA	THA	THA
xx012	HA	PA	CA	RA	TA	DA	HA	WA	SA	KA	DA	WA	JA	YA	NYA	MA	GA	BA	THA	WA
xx013	LA	PA	RA	RA	KA	DA	TA	SA	NA	LA	PA	DHA	MA	TA	NYA	MA	GA	HA	THA	NGA
xx014	HA	NA	CA	RA	LA	DA	TA	SA	WA	HA	PA	DHA	JA	YA	BA	MA	GA	BA	THA	NGA
xx015	YA	WA	CA	RA	KA	DA	TA	DHA	WA	HA	PA	DHA	JA	YA	NYA	WA	GA	NA	THA	NGA
xx016	LA	NA	CA	RA	DHA	DA	TA	SA	SA	HA	PA	WA	JA	YA	BA	MA	GA	BA	THA	KA
xx017	LA	MA	DHA	RA	KA	DA	TA	KA	JA	YA	SA	DHA	JA	YA	PA	MA	GA	SA	THA	DHA
xx018	HA	NA	MA	RA	LA	SA	TA	PA	JA	HA	PA	DHA	MA	YA	GA	MA	GA	BA	THA	NGA
xx019	LA	SA	CA	RA	KA	NA	YA	PA	PA	BA	PA	NA	MA	TA	SA	MA	GA	PA	THA	NGA
xx020	TA	NA	CA	RA	THA	DA	CA	SA	WA	YA	PA	WA	JA	YA	NYA	MA	GA	NYA	THA	NGA

Based on Table 2, the results of the HCC Model A test are transformed into a confusion matrix, making it easier to find the value of classification accuracy in testing with the HCC Model A method. The Confusion Matrix is shown in Table 3.

Table 3. Confusion matrix HCC model A

		RIIL																		
	Letter	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
	HA	6	0	0	0	1	0	3	0	0	7	0	3	0	0	3	0	0	1	0	0
	NA	0	7	0	0	0	2	0	1	1	0	3	1	1	0	0	0	0	2	0	0
	CA	0	1	10	0	0	0	1	0	2	0	0	3	0	0	0	1	0	0	1	1
P	RA	0	0	1	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
R	KA	0	0	0	0	10	0	1	3	0	2	0	0	0	0	0	0	0	0	0	2
E	DA	0	1	0	0	0	14	0	2	0	0	1	0	0	0	0	0	1	0	0	0
	TA	2	0	0	0	2	0	13	0	0	3	0	0	0	3	0	0	0	0	1	0
D	SA	1	3	0	0	1	2	1	8	2	1	1	0	0	0	1	0	0	1	0	0
I	WA	0	1	1	0	0	0	0	1	10	0	0	4	1	0	0	3	0	0	0	1
	LA	8	0	0	0	4	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
C	PA	0	3	0	0	0	2	0	3	1	0	13	0	0	0	1	0	1	2	0	0
T	DHA	0	0	2	0	1	0	0	2	0	0	0	9	0	0	0	0	0	0	0	1
I	JA	0	0	2	0	0	0	0	0	4	0	0	0	14	0	0	2	1	0	1	0
	YA	2	0	0	0	0	0	1	0	0	3	0	0	0	17	0	0	0	0	0	0
O	NYA	1	3	0	0	0	0	0	0	0	0	2	0	0	0	8	0	0	4	0	0
N	MA	0	1	2	0	0	0	0	0	0	0	0	0	4	0	0	14	0	0	0	0
	GA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	17	0	0	0
	BA	0	0	1	0	0	0	0	0	0	1	0	0	0	0	3	0	0	7	0	0
	THA	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	16	1
	NGA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	3	1	13

Based on the confusion table, the accuracy will be obtained by the following calculations:

$$Accuracy\ (\%) = \frac{6+7+10+20+10+14+13+8+10+3+13+9+14+17+8+14+17+7+16+13}{400} \times 100\%$$

$$Accuracy\ (\%) = \frac{229}{400} x\ 100\% = 57\%$$

According to the calculation results, the accuracy obtained is 57%.

3.2. Results of HCC Feature Extraction Method Model B

The next stage is to repeat the process of extracting HCC features from the intact image which is divided into 4 parts so as to obtain a number of 32 HCC feature parameters, then the training and testing process is carried out again with the same data, namely 1600 training image data and 400 test image data. The test results are presented in Table 4.

From the table the results of the HCC Model B test are transformed into a confusion matrix, making it easier to find the value of classification accuracy in testing with the HCC Model B method. The Confusion Matrix for Model B is presented in Table 5.

Table 4. Test results of HCC model B

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	BA	PA	CA	RA	KA	DA	TA	SA	WA	LA	SA	HA	JA	YA	NYA	MA	WA	NGA	THA	NGA
xx002	BA	THA	CA	CA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	NYA	NYA	MA	GA	HA	THA	NGA
xx003	HA	KA	CA	RA	KA	DA	TA	SA	CA	LA	PA	HA	JA	NYA	BA	MA	THA	BA	THA	NGA
xx004	KA	NA	WA	RA	KA	DA	TA	SA	WA	LA	PA	CA	LA	YA	NYA	MA	HA	BA	THA	NGA
xx005	HA	KA	CA	RA	KA	DA	TA	SA	WA	SA	THA	DHA	JA	NYA	NYA	MA	GA	BA	THA	NGA
xx006	HA	KA	RA	RA	KA	DA	JA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	HA	NGA	NGA
xx007	HA	NA	CA	CA	BA	DA	TA	SA	CA	LA	PA	DHA	JA	YA	NYA	MA	GA	NYA	THA	NGA
xx008	HA	KA	CA	RA	NYA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx009	HA	NA	CA	RA	NYA	PA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	HA	THA	NGA
xx010	TA	NA	CA	GA	KA	PA	TA	NA	WA	LA	DA	DHA	JA	YA	JA	MA	GA	HA	MA	MA
xx011	HA	NA	DHA	RA	KA	DA	DA	SA	PA	LA	THA	DHA	JA	LA	NYA	MA	GA	DHA	THA	NGA
xx012	HA	NA	CA	RA	KA	PA	TA	SA	WA	LA	THA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx013	BA	NA	CA	RA	KA	DA	BA	DA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	PA	NGA
xx014	HA	NA	CA	RA	LA	DA	TA	SA	WA	LA	PA	DHA	JA	TA	NYA	MA	GA	BA	THA	NGA
xx015	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	WA	JA	YA	NYA	MA	GA	BA	THA	CA
xx016	JA	NA	CA	RA	NA	DA	HA	NGA	PA	LA	PA	DHA	JA	YA	TA	MA	GA	BA	THA	NGA
xx017	JA	NA	CA	RA	TA	DA	CA	SA	WA	PA	PA	DHA	JA	YA	HA	MA	RA	MA	THA	NGA
xx018	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	THA	DHA	JA	YA	TA	MA	GA	BA	NGA	NGA
xx019	HA	YA	CA	RA	KA	NA	KA	SA	WA	DA	PA	DHA	JA	YA	NGA	MA	RA	BA	THA	NGA
xx020	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	HA	NYA	MA	GA	BA	PA	NGA

Table 5. Confusion matrix HCC model B

		RIIL																		
	Letter	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
PREDICTION	HA	13	0	0	0	0	0	1	0	0	0	0	2	0	1	1	0	1	4	0	0
	NA	0	13	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	CA	0	0	17	2	0	0	1	0	2	0	0	1	0	0	0	0	0	0	0	1
	RA	0	0	1	17	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	KA	1	4	0	0	14	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	DA	0	0	0	0	0	16	1	1	0	1	1	0	0	0	0	0	0	0	0	0
	TA	1	0	0	0	1	0	14	0	0	0	0	0	0	1	2	0	0	0	0	0
	SA	0	0	0	0	0	0	0	17	0	1	1	0	0	0	0	0	0	0	0	0
	WA	0	0	1	0	0	0	0	0	16	0	0	1	0	0	0	0	1	0	0	0
	LA	0	0	0	0	1	0	0	0	0	17	0	0	1	1	0	0	0	0	0	0
	PA	0	1	0	0	0	3	0	0	2	1	14	0	0	0	0	0	0	0	2	0
	DHA	0	0	1	0	0	0	0	0	0	0	0	16	0	0	0	0	0	1	0	0
	JA	2	0	0	0	0	0	1	0	0	0	0	0	19	0	1	0	0	0	0	0
	YA	0	1	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0
	NYA	0	0	0	0	2	0	0	0	0	0	0	0	0	3	14	0	0	1	0	0
	MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	1	1	1
	GA	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0
	BA	3	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	12	0	0
	THA	0	1	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1	0	15	0
	NGA	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	2	18

The confusion matrix Model B was used to calculate the accuracy. Accuracy is calculated in a manner as shown below:

$$Accuracy\ (\%) = \frac{13+13+17+17+14+16+14+17+16+17+14+16+19+14+14+20+15+12+15+18}{400} \times 100\%$$

$$Accuracy\ (\%) = \frac{311}{400} \times\ 100\% = 78\%$$

According to the calculation results, the accuracy obtained is 78%.

3.3. Results of HCC Feature Extraction Method Model C

The next step is to repeat the HCC feature extraction process from the intact image which is divided into 9 parts so that 72 HCC feature parameters are obtained. From the HCC feature parameters, the training and testing process was carried out again with the same data, namely 1600 image data to train and 400 image data to test it. The results of this test are presented in the form of Table 6.

The results presented in Table 6 are then transformed into a confusion matrix, making it easier to find the value of classification accuracy in testing with the HCC Model C Method. The Confusion Matrix resulting from the transformation is presented in Table 7.

Table 6. Test results of HCC model C

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	LA	DHA	JA	YA	NYA	MA	GA	HA	THA	NGA
xx002	BA	NA	CA	RA	KA	DA	TA	SA	WA	LA	HA	DHA	JA	YA	NYA	MA	GA	HA	THA	THA
xx003	BA	DA	CA	RA	KA	DA	TA	PA	WA	LA	PA	DHA	JA	NYA	NYA	MA	GA	BA	THA	NGA
xx004	DHA	NA	CA	RA	KA	JA	TA	SA	NGA	LA	PA	DHA	JA	YA	NYA	SA	GA	BA	THA	NGA
xx005	TA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	THA	YA	NYA	PA	GA	HA	THA	NGA
xx006	NYA	NA	CA	RA	KA	RA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	RA	BA	THA	NGA
xx007	HA	NA	LA	RA	NA	DA	HA	SA	WA	LA	HA	DHA	JA	YA	NYA	MA	GA	NYA	THA	NGA
xx008	HA	DA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	DHA	GA	BA	NGA	NGA
xx009	HA	NA	CA	RA	KA	SA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	TA	GA	BA	THA	NGA
xx010	PA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx011	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	KA	NYA	MA	RA	BA	THA	NGA
xx012	HA	NA	CA	RA	KA	DA	PA	SA	NGA	LA	THA	LA	JA	YA	NYA	MA	GA	HA	THA	GA
xx013	HA	NA	SA	RA	KA	DA	TA	SA	JA	LA	PA	DHA	JA	GA	NYA	MA	GA	BA	THA	PA
xx014	SA	NA	CA	RA	KA	DA	TA	SA	THA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx015	HA	NA	SA	RA	KA	NA	TA	SA	TA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx016	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	THA	DHA	JA	YA	NYA	MA	GA	BA	NGA	NA
xx017	HA	NA	CA	RA	KA	DA	TA	SA	WA	BA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	PA
xx018	HA	NA	NA	RA	TA	DA	TA	SA	WA	GA	PA	DHA	JA	NYA	BA	MA	RA	BA	THA	THA
xx019	HA	NA	CA	RA	KA	THA	KA	SA	WA	BA	JA	DHA	JA	NYA	NYA	MA	GA	BA	THA	NGA
xx020	HA	NA	CA	RA	KA	NA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA

Table 7. Confusion matrix HCC model C

		RIIL																		
	Letter	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
PREDICTION	HA	13	0	0	0	0	0	2	0	0	0	2	0	0	0	0	0	0	4	0	0
	NA	0	18	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	CA	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	RA	0	0	0	20	0	1	0	0	0	0	0	0	0	0	0	0	3	0	0	0
	KA	0	0	0	0	18	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
	DA	0	2	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	TA	1	0	0	0	1	0	16	0	1	0	0	0	0	0	0	1	0	0	0	0
	SA	1	0	2	0	0	1	0	19	0	0	0	0	0	0	0	1	0	0	0	0
	WA	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0
	LA	0	0	1	0	0	0	0	0	0	17	1	1	0	0	0	0	0	0	0	0
	PA	1	0	0	0	0	0	1	1	0	0	14	0	0	0	0	1	0	0	0	2
	DHA	1	0	0	0	0	0	0	0	0	0	0	19	0	0	0	1	0	0	0	0
	JA	0	0	0	0	0	1	0	0	1	0	1	0	19	0	0	0	0	0	0	0
	YA	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0
	NYA	1	0	0	0	0	0	0	0	0	0	0	0	0	3	19	0	0	1	0	0
	MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0
	GA	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	17	0	0	1
	BA	2	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	15	0	0
	THA	0	0	0	0	0	1	0	0	1	0	2	0	1	0	0	0	0	0	18	3
	NGA	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	13

The accuracy of the transformation whose result is the Confusion Matrix presented in Table 7 can be calculated from the following formula:

$$Accuracy\ (\%) = \frac{13 + 18 + 16 + 20 + 18 + 14 + 16 + 19 + 15 + 17 + 14 + 19 + 19 + 15 + 19 + 16 + 17 + 15 + 18 + 13}{400} \times\ 100\%$$

$$Accuracy\ (\%) = \frac{331}{400} \times\ 100\% = 83\%$$

According to the calculation results, the accuracy obtained is 83%.

3.4. Results of HCC feature extraction method model D

The next step is to repeat the HCC feature extraction process from the intact image which has been divided into 16 parts so that 128 HCC feature parameters are obtained. From the HCC feature parameters, the training and testing process was carried out again with the same data, namely 1600 image data to train and 400 image data to test it. The test results above are presented in the following tabulation of Table 8.

The results presented in Table 8 are then transformed into a confusion matrix to make it easier to find the value of classification accuracy in testing with the HCC Model D Method. The confusion matrix resulting from the transformation is presented in Table 9.

Table 8. Test results of HCC model D

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	HA	NA	CA	RA	HA	DA	TA	PA	WA	SA	WA	DHA	NA	YA	NYA	MA	GA	BA	THA	THA
xx002	HA	NA	CA	RA	KA	DA	KA	PA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx003	HA	NA	CA	RA	KA	DA	TA	SA	CA	LA	PA	NA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx004	HA	NA	CA	RA	KA	NA	TA	SA	GA	LA	SA	WA	JA	YA	NYA	MA	HA	BA	THA	NGA
xx005	KA	NA	CA	RA	KA	KA	TA	PA	DHA	LA	GA	DHA	JA	YA	NYA	THA	GA	BA	THA	BA
xx006	YA	NA	CA	RA	TA	NA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx007	HA	NA	CA	RA	KA	NA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	THA	GA	BA	THA	THA
xx008	HA	NA	DHA	RA	HA	SA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx009	HA	NA	NA	RA	BA	DA	TA	SA	WA	LA	DA	CA	JA	YA	NYA	MA	GA	BA	NGA	THA
xx010	HA	NA	CA	RA	TA	DA	TA	SA	CA	BA	DA	CA	JA	YA	NYA	MA	BA	BA	NGA	NGA
xx011	HA	NA	WA	RA	TA	PA	TA	SA	WA	LA	DA	WA	JA	YA	NYA	MA	GA	GA	NGA	NGA
xx012	HA	NA	CA	RA	KA	DA	JA	SA	DHA	LA	PA	CA	JA	YA	NYA	MA	GA	BA	THA	THA
xx013	HA	NA	CA	RA	TA	DA	TA	PA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	NYA	THA	NGA
xx014	HA	NA	WA	RA	NA	NA	KA	SA	NA	LA	PA	DHA	JA	HA	NYA	MA	GA	BA	THA	THA
xx015	HA	NA	CA	RA	KA	NA	TA	SA	CA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	TA
xx016	HA	NA	CA	GA	KA	DA	TA	SA	WA	LA	DA	DHA	JA	YA	NYA	MA	GA	HA	THA	NGA
xx017	PA	NA	WA	RA	HA	DA	WA	SA	WA	LA	DA	DHA	JA	YA	YA	MA	GA	BA	THA	THA
xx018	HA	NA	CA	RA	TA	DA	TA	SA	WA	LA	PA	CA	KA	YA	NYA	MA	GA	BA	THA	THA
xx019	YA	NA	DHA	RA	KA	DA	TA	SA	WA	LA	GA	DHA	HA	YA	NYA	MA	GA	BA	THA	NGA
xx020	LA	NA	DHA	RA	YA	NA	TA	SA	WA	LA	PA	DHA	JA	HA	NYA	MA	GA	BA	THA	NGA

Table 9. Confusion matrix HCC model D

		RIIL																		
	Letter	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
	HA	15	0	0	0	3	0	3	0	0	0	0	0	1	2	0	0	1	1	0	0
	NA	0	20	1	0	1	6	0	0	1	0	0	1	1	0	0	0	0	0	0	0
P	CA	0	0	13	0	0	0	0	0	3	0	0	4	0	0	0	0	0	0	0	0
R	RA	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
E	KA	1	0	0	0	9	1	2	0	0	0	0	0	1	0	0	0	0	0	0	0
D	DA	0	0	0	0	0	11	0	0	0	0	5	0	0	0	0	0	0	0	0	0
I	TA	0	0	0	0	5	0	13	0	0	0	0	0	0	0	0	0	0	0	0	1
C	SA	0	0	0	0	0	1	0	16	0	1	1	0	0	0	0	0	0	0	0	0
T	WA	0	0	3	0	0	0	1	0	13	0	1	2	0	0	0	0	0	0	0	0
I	LA	1	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0
O	PA	1	0	0	0	0	1	0	4	0	0	11	0	0	0	0	0	0	0	0	0
N	DHA	0	0	3	0	0	0	0	0	2	0	0	13	0	0	0	0	0	0	0	0
	JA	0	0	0	0	0	0	1	0	0	0	0	0	17	0	0	0	0	0	0	0
	YA	2	0	0	0	1	0	0	0	0	0	0	0	0	18	1	0	0	0	0	0
	NYA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0	0	1	0	0
	MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0
	GA	0	0	0	1	0	0	0	0	1	0	2	0	0	0	0	0	18	1	0	0
	BA	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	17	0	1
	THA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	17	10
	NGA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	8

The accuracy of the transformation whose result is the Confusion Matrix presented in Table 9 can be calculated from the following formula

$$Accuracy\ (\%) = \frac{15 + 20 + 13 + 19 + 9 + 11 + 13 + 16 + 13 + 18 + 11 + 13 + 17 + 18 + 19 + 18 + 18 + 17 + 17 + 8}{400} \times 100\%$$

$$Accuracy\ (\%) = \frac{330}{400} \times\ 100\% = 76\%$$

According to the calculation results, the accuracy obtained is 76%.

3.5. Results of Comparative Analysis of HCC Method

This research has carried out 4 (four) models of HCC feature extraction. The recapitulation of the four models is presented in Table 10.

Table 10 shows the lowest accuracy obtained from the extraction of Model A, which is 57%. The highest accuracy was obtained from the Model C extraction, which was 83%. Models with intact images, namely those that are not divided, produce the lowest accuracy values compared to the divided images.

However, making the image into a larger number does not always increase the accuracy value. This can be seen from the division of the image into 16 parts only produces an accuracy value of 76%. This result is smaller than the image which is divided into 9 parts with an accuracy value of 83%.

Table 10. HCC model feature extraction testing accuracy

HCC Feature Extraction Model	Number of Chain Code Segments	Number of HCC Feature Parameters	Accuracy
A	1	8	57%
B	4	32	78%
C	9	72	83%
D	16	128	76%

4. Conclusion

This research studies the application of the Histogram Chain Code (HCC) method as an image processing method for Old Javanese letters. The number of image parameters is multiplied by dividing the whole image into several parts with a total of 4, 9, and 16 divisions. The division of the image with these 3 divisions is compared with the whole image (not divided). The accuracy of the computer's ability to recognize the image was compared experimentally between the divided image and the whole image. From the experiment, the HCC model which divides the image into 3x3 pieces or divides it into 9 parts which obtains 72 feature parameters produces the highest accuracy compared to images that are not divided or divided into other numbers. Characteristics of the Javanese script which has special characteristics between one letter and another causes the division of the image into several parts to have the best number of image divisions compared to the division with other numbers. This study also found that dividing the image into more parts does not always increase the level of accuracy compared to smaller divisions. Javanese script letters that have unique shapes such as Ra, Sa, La, Ja, and Nya have a good level of accuracy compared to scripts that look similar to each other. This requires improvement of methods that can capture the special characteristics of the Javanese script to increase its accuracy. This study also suggests that further research should compare other chain code methods, such as Differential Chain Code and Vertex Chain Code.

References

[1] H. Surjono, “Multimedia Pembelajaran Interaktif: Konsep dan Pengembangan,” UNY Press, 2017.

[2] “PERGUB Prov. DIY No. 64 Tahun 2013 tentang Mata Pelajaran Bahasa Jawa Sebagai Muatan Lokal Wajib di Sekolah atau Madrasah [JDIH BPK RI].” [Online]. Available: https://peraturan.bpk.go.id/Home/Details/21626. [Accessed: 30-May-2022].

[3] Isnaeni, “PENGEMBANGAN MEDIA PEMBELAJARAN INTERAKTIF AKSARA JAWA BERBASIS FLASH PADA MATA PELAJARAN BAHASA JAWA,” 2016. [Online]. Available: http://lib.unnes.ac.id/27938/. [Accessed: 16-Apr-2022].

[4] K. Prihatin, “PENGEMBANGAN MULTIMEDIA INTERAKTIF AKSARA JAWA UNTUK PEMBELAJARAN BAHASA JAWA SISWA KELAS V SD NEGERI SABDODADI KEYONGAN BANTUL - Lumbung Pustaka UNY,” 2015. [Online]. Available: https://eprints.uny.ac.id/17776/. [Accessed: 16-Apr-2022].

[5] H. Freeman and L. S. Davis, “A Corner-Finding Algorithm for Chain-Coded Curves,” IEEE Trans. Comput., vol. 26, no. 03, pp. 297–303, Mar. 1977, doi: 10.1109/TC.1977.1674825.

[6] S. Mawaddah and N. Suciati, “Pengenalan Karakter Tulisan Tangan Menggunakan Ekstraksi Fitur Bentuk Berbasis Chain Code,” J. Teknol. Inf. dan Ilmu Komput., vol. 7, no. 4, pp. 683–692, Aug. 2020, doi: 10.25126/JTIIK.2020742022.

[7] A.-B. M. Salem, A. A. Sewisy, and U. A. Elyan, “A Vertex Chain Code Approach for Image Recognition.”

[8] S. Arora, D. Bhattacharjee, M. Nasipuri, D. K. Basu, and M. Kundu, “Application of Statistical Features in Handwritten Devnagari Character Recognition,” Jun. 2010, doi: 10.48550/arxiv.1006.5911.

[9] Y. Qian, W. Xichang, Z. Huaying, S. Zhen, and L. Jiang, "Recognition Method for Handwritten Digits Based on Improved Chain Code Histogram Feature," Proc. 3rd Int. Conf. Multimed. Technol., vol. 84, pp. 431–438, Nov. 2013, doi: 10.2991/ICMT-13.2013.53.

[10] R. Vickery, "8 Metrics to Measure Classification Performance | by Rebecca Vickery | Towards Data Science," 2021.